戴玉数据圈 1.8
城市数据团：从数据工作者到自媒体“主编”
——跨界的惊喜、心得与迷惑

“戴玉数据圈”是资深数据新闻人戴玉创办的数据新闻群，里面汇聚了各类数据报告发布方、数据媒体主编和记者、高校数据新闻教师、可视化人员、数据行业人员等等，旨在搭建切实推动数据新闻行业交流与合作的平台。

【数据圈那些人】系列沙龙，主要用来初步介绍群内伙伴们各自的工作，形成初步了解。

分享人：汤舸 “城市数据团”、“脉策数据”联合创始人，于同济大学和伦敦大学学院主修建筑和城市规划以及城市更新

分享时间：2017.03.03 下午 16:30

嘉宾分享环节

“城市数据团”的 slogan 是用数据阅读城市，主要做一些基于城市的数据研究工作，有些文章在互联网上广为流传，我们利用各种各样的数据源，针对城市的热点问题，进行深入挖掘和分析，并从专业的研究转化为大众化的文本。我们团队本身是做城市研究和数据分析的，并不是媒体工作者。由于做了『城市数据团』这个公众号，误打误撞地进入到了媒体行业，并且是进入到了最火的数据媒体行业。

但今天想跟大家分享的并不是媒体创作的心得，而是作为一个媒体行业外的人，针对这个领域的一些观察和思考。我的这次分享就从城市数据团的团队构成谈起吧。

1. 数据团的自媒体工作流程

数据团的工作流程可以划分为：数据采集、数据清洗、数据建模与分析、可视化、研究成果制作五个部分。与一般的长周期学术研究成果不同，我们对于一个主题的研究产出时间保持在 1–2 周的短周期。而在这个周期当中，团队成员的 90% 时间是花费在数据整理、清洗、验证假设、调整模型这四个阶段的。10% 的时间花费在可视化制作以及研究文章撰写方面。

从这个角度来说，数据团自媒体团队的构成可以分为数据工作组和成果工作组两个组合。其中数据工作组的总人数约为 1.5 人，成果工作组的总人数约为 1 个人。简而言之，两个组加起来的总人数大约是 2 个人。而这样的构成对个人的要求就会变得很高，必须又懂数据、又懂研究、又懂画图，又懂写文字。

很显然，这就会是一个很脆弱的结构，因为把大部分的工作集中在一个人身上时，一旦某个人临时有点别的事，自媒体的工作自然就会受到严重的影响。但幸运的是，这样具有综合技能的同事，在数据团里还有好多个，只不过都在处理自媒体以外的工作，但在特殊时期都可以临时来帮忙，以保证自媒体产出的稳定。（但其实有时候我们也很随意，不稳定就不稳定了也没什么。）

2. 团队运营模式

那么问题来了，我们为什么在组织上采用这种脆弱的结构，而不是采用工业化的稳定结构，把数据采集、清洗、分析、建模、可视化、文章等工作分成不同环节，分给不同的同事做呢？大概有以下几个原因：

1. 工业化的体系模式只在规模化生产时才能产生其优势，在自媒体产量不高的情况下，采用个体化生产的模式会比工业化体系更轻松。

2. 创作本身是一个高度需要凝聚力和试错成本很高的工作，一个人单独负责创作的全流程，往往比多人协作产生的效果更好。

3. 自媒体挣钱太困难，必须控制好成本才能够持续稳定地运营下去。

而这也并不是一开始我们就想明白的问题。从创立自媒体起始，我们平均每个月都会有一篇刷屏朋友圈的文章出现，每次这种文章出现后，我们就会收到如雪片般飞来的关注、留言、合作意向、投资意向、项目咨询等等，从普通读者到业界大佬，数不胜数。最开始，对于我们这些没见过世面的研究者而言，这种阵势是很容易让人迷失的。会让人产生“我们太牛逼了，赶紧招人投钱扩大吧”这样的幻觉。事实上呢？自媒体由依托内容价值的商业之路（转电商的不算）至今仍未被证实，更别说自媒体了，连传统媒体也都一样。何况数据类自媒体相比传统自媒体是更烧钱的一个游戏。所谓更烧钱，主要来自于行业间人力成本的差异。

总体而言，媒体行业从业者的平均收入远远低于数据科技行业从业者的平均收入。而数据类自媒体又高度依赖后者。使得数据媒体的人力成本显得十分不健康。因此，我们大概明白了一件事：数据类自媒体（非体制内的）想要活下来，就必须要赚到比一般媒体更多的钱才行。其实，这也是所有体制外的自媒体的最核心要务：赚钱。

3. 关于商业价值

有人说做影响力也很重要，影响力大了，自然就有商业价值。其实我个人是很反对（或者说反感吧）这个观点的。短期来看，影响力 = 商业价值；但长期来看，正是这种观点害死了传统媒体。传统媒体行业为什么衰落了？为什么大量的人才从传统媒体行业中流失？一般人们会认为是媒体渠道分散、新媒体崛起等外部原因。

其实并不完全是这样，传统媒体行业的衰落，其本质原因在于传统媒体的商业模式的缺陷。传统媒体的商业模式是建立在广告收费而非内容收费上的。这是一个特别不健康的商业模式。举个简单的例子，为什么电影视频等新娱乐方式崛起了，而小说为什么还没有衰落呢？因为小说是靠内容收费而非广告收费的。同样都是面向 C 端创作有价值的内容，但媒体行业的商业模式是 to B 的，而小说的商业模式则是 to C 的。不向终端用户收费的商业模式在本质上都是不能长久的。正是这种商业模式的不健康，使得传统媒体失去了人才上的竞争力，从而失去了行业竞争力。

作为一个媒体行业外的人和团队，我们给自媒体的发展设定了三条原则，也是基于以上的思考和考虑：①做自媒体一定要赚钱②但永远不做广告③要通过内容本身的价值赚钱。我们围绕着这三条原则，做了很多商业化的尝试。

最近，我们做了『数据技能在线教育』这个模块，感觉还不错。

数据链接： https://datateam.ke.qq.com

从页面的报名数据和客单价就可以看到，这还是一个不错的项目。基本营收可以 cover 整个数据团自媒体部门的支出。有了商业化的支持，我们才可以更任性地做一些有关这个城市的深入研究、更任性地继续做自媒体研究、更任性地去不计成本地做一些自己想做的东西，更任性的去用数据发现城市的秘密。没有健康商业化模式的媒体是无法持续的，这是我们作为外行人进入到自媒体行业后，最深刻的感受。都是一些跟自媒体创作无关的事，但我个人觉得，这些事比自媒体创作更值得思考。否则今天看上去热火朝天的自媒体行业，接下来明天很可能就是日渐式微的传统媒体。

交流提问环节

（遵照发言嘉宾的要求，内容略有删节）

主持人戴玉： 汤老师是数据行业的人，平时也很难“逮到”，其实我自己有特别多的问题想问你。第一个就是，数据团生产一篇稿件，大概的成本（包括人力、办公）有多少？是否计算过？换句话说，如果有媒体想直接订制购买你们的一篇文章，成本如何？

汤舸： 虽然部门只有 2 个人，但也是因为镶嵌在整个数据团内部才能够 work，因此更多的都是共享资源。我们自己核算过，一年整个自媒体的投入在 300-400 万左右。换算到单篇文章差不多是 10-15 万这样的成本规模。

主持人戴玉： 这么高！你们是怎么承担这笔费用的呢？

汤舸： 我们有赚钱的主业（笑）。

学生 ANNA： 现在网络上数据都亦真亦假，那要怎么才能正确地采集最真实的数据呢？

汤舸： 很难说最真实吧，只能说逼近真实。逼近真实的方法就是多元数据的比对、清洗和校验。比如说上海的人口问题，我们就需要拿来和住房增长以及轨道交通流量增长等数据放在一起校验，才比较靠谱。

主持人戴玉： 我还有一个问题，平时我们数据新闻行业，其实很想招一些数据分析师进来写文章，但是相当难招到。要么就是觉得薪水低，来看看就走了；要么就是留不住。所以特别想知道，在数据分析师眼里，跟媒体岗位竞争的其他岗位和其他诱惑有哪些？如果分析师来写报道，吸引他的又是什么？

汤舸： 这个问题好啊，不过我感觉这个问题也没那么复杂。写报道这件事，肯定是不会能够吸引数据工作者的。那么怎么办呢？给高工资呗，很简单啊。不然为什么数据自媒体成本那么高。【听到这里，数据圈友纷纷发表情（捂脸）（哭）】给钱之后谈理想才是真正的理想主义者，不给钱谈理想是纯流氓。我国自媒体该阶段的主要问题是解决赚钱问题。

财新黄晨： 解决不了盈利问题的高薪都是泡沫。

百度 ECharts_ 祖明： 汤老师能谈谈今日头条这样的 feed 产品对于自媒体的价值和影响吗？包括商业收入层面会有什么贡献吗？

汤舸： 我还是之前的观点：自媒体唯一健康的商业模式就是通过内容向终端用户收费。通过广告都是很难持久的。因为广告依赖的是分发能力，分发能力只是媒体的水流，源头还是内容创作。

财新黄晨： to C 的内容收费，在中国，很难。

主持人戴玉： 国内外媒体好像都建立过付费墙，付费阅读，后来效果好像不太理想啊。

百度 ECharts_ 祖明： 面向终端用户是说只 to c 吗？现在有 to b 的模式吗？比如某家头条号找您买内容，然后按访问量计费啥的？

汤舸： 只要是为内容付费的，to B 也是健康的。所以您说的这种模式我个人认为是挺好的。

小尤 - 上海观察： 我们就建过付费阅读墙，总体来说一般化新闻没有付费必要，收费内容要么是娱乐需求，要么是学习或者业务需求。前者是 A 片，后者就是各种情报，针对个人用户就是学习参考资料。没有人会为新闻付钱，分析性的新闻也不行。海外的报纸可以。我付钱看过一年纽约时报和 FT，说实话这个内容拿到中国来一样失败，FT 号称有很多分析的，但话题还是大众型的，不是行业型、专业型的。媒体的切入点就在于你不是一个专业的切入，数据分析的角度也是媒体的角度。

中北大学任占文老师：国人的版权意识和消费习惯不一样。

澎湃王昀：很简单的一个问题，数据分析如果真的能够揭示某个有商业价值的奥妙，为什么要所有公众买单呢？直接卖给相关企业不好么，如果相关企业不肯买，就意味着则个数据分析没有新的发现，更不要拿来忽悠公众了。靠公众付费是不能支撑靠谱的分析的，如果有这样的先例，请让我看到。

汤舸：数据分析和新闻一样，最后还是要终端用户付费才能持续。但无非终端用户是 to B 或者 to C 的区别。大众需要的不见得是分析。

澎湃王昀：结论就是大众不配享用分析？这本身就与媒体起初“开启民智”的指向相悖了。显然大众是不愿付费的。

汤舸：不是不配享用啊，而是只有付费享用才能够使得高质量内容源可持续地被生产。说到大众付费的问题，我的观点是这样的：如果大家不愿意为好电影付费，那么活该大家看烂片。道理简单得很。

小尤 - 上海观察：我前段时间接触了一些 EMBA 的班，他们请一些课外讲师，我觉得好 low，比如讲数据的时候。但是真的很受欢迎，一天的课程六七万。人家是完全从老板的需求出发。

主持人戴玉：我感觉在媒体做数据分析还是跟智库不一样，首先要解决的问题就不一样。媒体的数据分析首先要用来回应读者期待，就是脑中有个意识是“受众想看什么，受众迷惑什么”。但看真正的分析报告的时候，一般都是以问题本身的重要性排序。某个问题本身最重要的是什么，就去分析什么。

沈杭珍 - 我在现场 app：新闻本来就是信息的一种。去中介化后，新闻内容最终要和信息一样对待。没有价值的新闻，冠上新闻二字还是没有价值。说的价值，就是满足不同人群的不同需求。

百度 ECharts_ 祖明：我看汤老师你们的作品里的图表都很漂亮，一般都是用什么工具做的呢？

汤舸：一般都是 excel（捂脸）。

学生嗒嗒：谈到秘密，想请问一下你们数据公开涉及到的公共部分或企业或其它的秘密，这些秘密的公开是否曾造成误伤，或者其它不未曾料到的误伤，又怎么处理？

汤舸：我们这点特别谨慎，基本不针对私人机构做任何评判，只涉及公共部分，公共部分就是公域，大家都可以探讨的。觉得我们不对的就说呗，我们接受批评。我记得我们研究双十一假打折的时候把所有商家的名称都打了码（捂脸），就怕误伤。

百度 ECharts_ 祖明：商业模式这个我觉得因人而异，不同的人有不同的赚钱方式，适合自己团队，达到自己团队目标即可，肯定不是大一统的。不适合别人的未必不适合自己，适合自己的也未必所有公司都通用。

汤舸：那倒是的，非常赞同。

小尤 - 上海观察：同意。现在赚钱的方式很多，关键是想好给谁看。想好了这点，一般模式就很清晰。给大众看，付钱的就是广告主了，所以商业模式就是 to B。

汤舸：是啊，只不过 2c 服务 2b 赚钱这种模式的链条会比较长，相对来说会有一定的脆弱性，当然也没那么绝对。报纸不行了不就是个典型的例子么。

沈杭珍 - 我在现场 app：2B 和 2C 会有个人员配置和自由度的问题。2B 的人员配置，要考虑公关人才。2C，要考虑传播人才。

丁利 -199IT：有一个观点不是很认同，影响力，其实就是商业价值。因为影响力即用户，没有用户一切空谈。至于有的人直接通过向 C 端直接收钱，还有是媒体向 B 端收钱，模式不一样。在这 C 端与 B 端这二方向上，其实都有应该相互借鉴与学习。

汤舸：嗯，这个话题特别有趣，可以讨论很久。我的观点是抛砖引玉啦。

丁利 -199IT：说到今日头条，其实，今日头条的广告营收去年好像已经达到 60 亿，如果今日头条做培训，估计也就是 6 个亿的收入而已。我只是假设一下向用户直接收钱的模式。直接向用户收钱的风险很大，因为这是在洗用户。直接向 B 端收费能在媒体里流行百年，是有道理的。

吕桂捷 -CNNIC：说到数据机构的媒体化，我也想要补充一些，给大家做个思路的拓展，因为我们 CNNIC 一直都是“去媒体化”，不去探讨热点，不用话题性语言，拒绝话题性机构合作，还有就是不做排行榜。以上去媒体化，绝大多数都是为了保持报告、数据的权威性。而正因为权威性强，所以有些机构来付费找我们做研究（当然做完以后也不会对外发布，基本上就是机构自己参考，或者上报给领导机构看，体现本区域、本部门信息化或者相关方面工作成果，再看看还有哪些可以提升的空间）。我们 2C 是完全免费，公益报告。前段时间跟腾讯的企鹅智库分析师聊了一下，他说腾讯的行业报告很快会走向收费，问我们会不会收费，我们以后也不会。我们的报告还比较粗，主要还是发调查数据，其实也有细的研究但是因为种种原因没有写进去，所以很多时候都是作为资料性的参考，对趋势性的东西表露不多（其实研究还是挺多的，就是不能在报告讲太多）。所以很多时候，商业性报告作为更充分、更细致、更有前瞻性的资料也很有很大价值，各自分工不同。

主持人戴玉：圈友们，我们本次沙龙持续了快一个半小时了，在汤老师充满个性的观点表达下，大家表达了不少真知灼见，也谢谢城市数据团的 @ 汤舸 和各位的参与！媒

（整理人 / 李安娜 高赫）